Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 362–366, January 1984
Biochemistry

Avian sarcoma virus *gag* and *env* gene structural protein precursors contain a common amino-terminal sequence

(hybrid-selection/*in vitro* translation/amino-terminal peptide analysis/mung bean nuclease mapping)

THOMAS A. FICHT, LUNG-JI CHANG, AND C. MARTIN STOLTZFUS

Department of Microbiology, University of Iowa, Iowa City, IA 52242

Communicated by Aaron J. Shatkin, September 22, 1983

ABSTRACT **The initiation site for translation of the avian sarcoma virus glycoprotein precursor, $Pr63^{env}$, has been determined by analyzing the amino-terminal peptides of $Pr63^{env}$ and the polyprotein precursor $Pr76^{gag}$ encoded by the viral *gag* gene. The acceptor splice junction used to form the *env* gene mRNA has also been identified. Hybrid-selected virus-specific mRNAs were translated *in vitro* in the presence of either L-[^{35}S]methionine to label at every methionine residue or L-[^{35}S]methionine-$tRNA_i^{Met}$ to label specifically at the amino-terminal methionine residues. Tryptic peptide maps of $Pr63^{env}$ labeled at every methionine residue contain all of the peptides, plus one additional peptide, present in the map of $Pr57^{env}$, a nonglycosylated *env*-encoded polypeptide of molecular weight 57,000 immunoprecipitated from tunicamycin-treated cells. Specific amino-terminal labeling of the *in vitro*-synthesized polypeptides showed that the peptide missing from $Pr57^{env}$ corresponds to the amino-terminal tryptic peptide of $Pr63^{env}$, which is removed *in vivo* as part of the amino-terminal signal peptide. Comparison of the amino-terminal tryptic peptides of $Pr63^{env}$ and $Pr76^{gag}$ showed that they are identical. In contrast, the chymotryptic amino-terminal peptides of $Pr76^{gag}$ and $Pr63^{env}$ are not identical. The location of the acceptor-splice junction in the *env* mRNA of the Prague A strain of avian sarcoma virus was determined by mung bean nuclease mapping to be at nucleotide 5,078. Fusion of the *gag* and *env* gene sequences during splicing results in use of the same AUG codon to initiate synthesis of $Pr76^{gag}$ and $Pr63^{env}$. This sequence is contained within the 397-nucleotide 5′ terminal leader that is spliced to the body of the *env* mRNA. The possible significance of these results for the regulation of avian sarcoma virus synthesis and translation is discussed.**

Productive infection by the avian leukosis viruses or the transformation-defective avian sarcoma viruses (tdASV) requires the formation of an 8.4-kilobase (kb) RNA that serves either as genomic RNA or as mRNA for the 76-kilodalton (kDa) precursor ($Pr76^{gag}$) of the virion nonglycosylated structural proteins (1–3) and a 3.5-kb subgenomic spliced RNA that serves as mRNA for the glycoprotein precursor (4). There are two AUG codons from which the synthesis of the envelope proteins could be initiated. The sequences in the Prague C strain of Rous sarcoma virus (RSV) genome known to code for the amino-terminal sequence of gp85 are present immediately downstream from nucleotide 5,245 (using the convention in which nucleotide 1 is the 5′ terminus of the RSV genome; refs. 5 and 6). An AUG codon in phase with these sequences is present 191 bases upstream at nucleotide 5,054. Therefore, if initiation of the *env* gene product occurs at this site, then the primary *env* gene product would be a 63-kDa polypeptide product. The use of this AUG codon as an initiator implies that the acceptor splice junction lies in the sequences upstream of this initiation codon. Nucleotides 4,868–4,876 make an excellent potential consensus splice acceptor site for this possibiity (5, 7). As a second possibility, the initiator AUG may be contained in the 5′-terminal 397-base leader sequence that is spliced to the body of the *env* gene sequences (5, 8, 9). S1 nuclease mapping experiments suggest that the donor splice junction used to form the subgenomic mRNA of the Schmidt–Ruppin A strain of RSV (SRA) is 18 nucleotides downstream from the *gag* gene initiation codon (8, 9). Therefore, it is possible that the AUG codon used to initiate synthesis of $Pr76^{gag}$ might also be used to initiate synthesis of the envelope polypeptide precursor. A potential splice acceptor sequence that lies 24 nucleotides downstream from the AUG codon at nucleotide 5,054 may be used in this case (5, 6). The primary *env* gene product in this case would not be significantly different in size than that predicted by initiation at nucleotide 5,054. The two models can be distinguished by analysis of the amino-terminal sequences of the respective primary gene products. The first model—i.e., initiation of the *env* gene product at nucleotide 5,054—predicts that the amino-terminal sequences of $Pr76^{gag}$ and $Pr63^{env}$ are nonidentical. Initiation within the 5′-terminal leader sequences, on the other hand, predicts that the amino-terminal sequences of the two gene products are identical. In this study, we have distinguished these two possibilities by analyzing the amino-terminal peptides of the *in vitro* translation products of hybrid-selected virus-specific mRNA isolated from ASV-infected cells. We have also localized by mung bean nuclease mapping the splice acceptor site of the RSV *env* mRNA. Our results support the hypothesis that initiation of the *env* gene product occurs within the 5′ leader sequence and that the 5′ leader sequences are spliced to a consensus acceptor site downstream from the AUG at nucleotide 5,054 (5, 6).

MATERIALS AND METHODS

Cell Culture. Cultures of chicken embryo fibroblasts were prepared, maintained, and infected with the Bratislava 77 strain of tdASV as described (10).

Immunoprecipitation of Virus-Specific Proteins. Infected cells were incubated for 4 hr in low-methionine medium, labeled for 1 hr with L-[^{35}S]methionine (40 μCi/ml, 1500 Ci/mmol; 1 Ci = 37 GBq), and harvested, and the virus-specific proteins were immunoprecipitated as described (11–13). Tunicamycin (1 μg/ml; Sigma) was used to inhibit glycosylation (6). Cells were incubated in the presence of L-[^{35}S]methionine for 12 hr to label virion proteins and the virus was purified as described (10).

Purification of ASV-Specific mRNA. Infected cells were labeled with [^{3}H]uridine (5–10 μCi/ml) for 12–36 hr and the polysomes were isolated as described by Lee *et al.* (14). The

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: kb, kilobase(s); kDa, kilodalton(s); ASV, avian sarcoma virus; tdASV, transformation-defective ASV; RSV, Rous sarcoma virus; SRA, Schmidt–Ruppin A strain of RSV.

RNA was purified as described by Strohman *et al.* (15), and virus-specific mRNA was isolated by hybridization-selection essentially as described by Stoltzfus and Dane (11). This RNA was separated into size classes by sedimentation on 5–30% (vol/vol) glycerol gradients, and the amount of RNA in each fraction was determined from the specific radioactivity of the polysomal poly(A)-containing RNA.

***In Vitro* Synthesis of Virus-Specific Proteins.** Virus-specific RNA (2–8 μg/ml) was translated in rabbit reticulocyte lysates (Amersham) with L-[^{35}S]methionine (1,500 Ci/mmol, 2,000 μCi/ml) as described by Pelham and Jackson (16). L-[^{35}S]Methionine-labeled $tRNA_i^{Met}$ was prepared from calf liver tRNA (Boehringer Mannheim) as described by Dubnoff and Maitra (17) and was added at a final concentration of 400–800 μg/ml (500,000 cpm/μg) along with 500 μM methionine/40 μM dithiothreitol.

$NaDodSO_4$/Polyacrylamide Gel Electrophoresis. $NaDodSO_4$/polyacrylamide slab gels were prepared as described by Laemmli (18). Radiofluorograms of the gels were prepared as described by Bonner and Laskey (19) or by Chamberlain (20) and exposed to Kodak Blue brand film BB5.

Peptide Mapping. Tryptic and chymotryptic peptide maps were prepared as described by Pawson *et al.* (21).

RESULTS

Genomic 35S RNA and subgenomic 21S RNA were purified from the polysomes of cells infected with the Bratislava strain of tdASV as described above and in refs. 13, 15, 22, and 23. Translation directed by 35S RNA (8.4 kb) resulted in synthesis of $Pr76^{gag}$ (Fig. 1*A*, lanes 1–5). Immunoprecipitation and tryptic peptide analysis confirmed the identity of the major *gag* gene product and the polypeptide bands of higher mobility in lanes 1–5, which result from either premature termination or cleavage of $Pr76^{gag}$ (unpublished results). The precursor to the reverse transcriptase, $Pr180^{gagpol}$, was synthesized at reduced levels and was visible only after increased exposure of the autoradiogram shown (data not shown). Translation from 21S (3.5-kb) RNA produced a 63-kDa polypeptide (Fig. 1*A*, lanes 6–8) that is the presumptive precursor to the mature glycoproteins (24, 25). Confirmation of the identity of the 63-kDa polypeptide was obtained by immunoprecipitation using antiserum raised against the mature envelope glycoproteins (refs. 26–28; Fig. 1*B*, lane 6). The specificity of the antiserum is shown in Fig. 1*B* (lane 2) and indicates that only the mature glycoproteins are immunoprecipitated from detergent-disrupted virions. This same antiserum immunoprecipitates the 95-kDa glycosylated precursor (gPr95) to the mature glycoproteins from extracts of infected cells (refs. 29–32; Fig. 1*B*, lane 4) and the 57-kDa nonglycosylated precursor ($Pr57^{env}$) from extracts of cells treated with the glycosylation inhibitor tunicamycin (refs. 6, 33, and 34; Fig. 1*B*, lane 5). The apparent difference in the molecular weights of the *in vivo* nonglycosylated *env* product precursor and the *in vitro* translation product of the *env* gene (Fig. 1*B*, lanes 5 and 6) can be attributed to *in vivo* processing of the amino-terminal signal peptide. This proteolytic cleavage is not inhibited by tunicamycin. These data are consistent with previous reports indicating that the ASV envelope precursor has a signal peptide of approximately 62 amino acids (5, 6, 33).

The relationship between the mature envelope glycoproteins and the putative precursors was examined by tryptic peptide analysis of the L-[^{35}S]methionine-labeled proteins. A tryptic peptide map of a mixture of the mature envelope glycoproteins gp85 and gp37 that were immunoprecipitated from disrupted virions is shown in Fig. 2*A*. Four methionine-containing peptides were identified in the tryptic peptide map of gp37 and correspond to peptides 1–4 (unpublished results). A single additional peptide spot (peptide 5) distin-

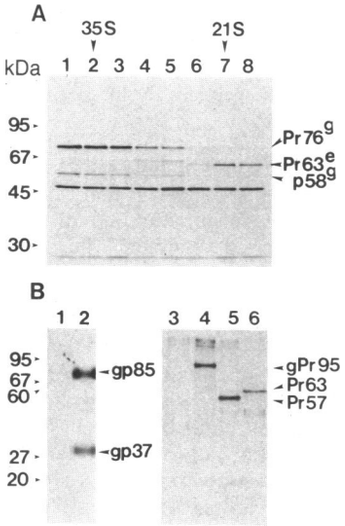

FIG. 1. *In vitro* translation of virus-specific gene products and immunoprecipitation of *in vivo*-synthesized glycoprotein precursors. (*A*) Virus-specific mRNA was purified by hybrid-selection and size fractionated on 5–30% (vol/vol) glycerol gradients. Aliquots from each fraction were translated, and the polypeptide products from 5-μl reaction mixtures were separated by electrophoresis on an 8.75% $NaDodSO_4$/polyacrylamide gel, and the gel was prepared for radiofluorography. A 48-kDa endogenous translation product is present in all the lanes. (*B*) Envelope glycoprotein antiserum was used to immunoprecipitate L-[^{35}S]methionine-labeled mature glycoproteins from disrupted virions or glycoprotein precursors from cell lysates. The polypeptide products were separated by electrophoresis on a 10% $NaDodSO_4$/polyacrylamide gel. Lanes: 1, normal rabbit serum with disrupted virions; 2, glycoprotein antiserum with disrupted virions; 3, normal rabbit serum with untreated cell lysate; 4, glycoprotein antiserum with untreated cell lysate; 5, glycoprotein antiserum with tunicamycin-treated cell lysate; 6, glycoprotein antiserum with 21S viral RNA *in vitro* translation product.

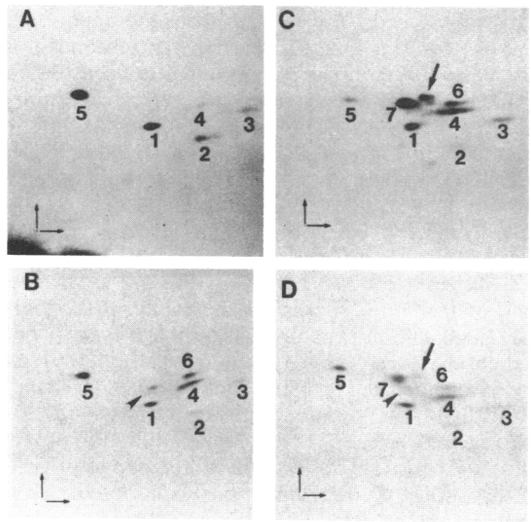

FIG. 2. Tryptic peptide maps of mature envelope glycoproteins and nonglycosylated *in vivo*- and *in vitro*-synthesized precursors. The proteins were isolated as described in the legend to Fig. 1 and excised from the dried gels. Oxidation, digestion with ketone-treated trypsin, and electrophoresis were carried out as described (21). Electrophoresis was from left to right and followed by ascending chromatography. (*A*) gp85 together with gp37 (2,500 cpm of each); (*B*) $Pr57^{env}$ (2,500 cpm); (*C*) $Pr63^{env}$ (5,000 cpm); (*D*) $Pr57^{env}$ together with $Pr63^{env}$ (2,500 cpm of each). Peptide maps were dipped in molten 2-methylnapthalene containing 0.4% (wt/vol) diphenyloxazole. Exposures were for 10–14 days using Kodak X-Omat X-AR5 film. The identities of peptides labeled by arrows are not clear and the presence of these peptides is variable.

guishable on the tryptic maps was obtained from a mixture of gp85 and gp37. This result is due to extensive glycosylation of gp85 (6), which produces a heterogeneous smear of the glycosylated peptides visible on longer exposure of the autoradiogram shown (data not shown). The peptide map of the nonglycosylated precursor (Pr57^{env}) contained the peptides that originate from sequences encoding gp37 at its carboxyl terminus (Fig 2*B*, spots 1–4), as well as two additional peptides that presumably originate from the amino-terminal region containing the gp85 sequences (refs. 6, 35, and 36; Fig. 2*B*, spots 5 and 6). An additional peptide present in the map of Pr57^{env} (Fig. 2*B*, arrowhead) was not assigned because its recovery varied and it was difficult to detect in the tryptic map of Pr63^{env}.

The peptide map of the primary translation product of the envelope gene, Pr63^{env}, contained all of the peptides present in the tryptic map of Pr57^{env}, as well as two additional peptide spots that were not prominent in the tryptic map of Pr57^{env}. One spot (Fig. 2*C*, spot 7) that is always present in the tryptic maps of Pr63^{env} and never observed in tryptic maps of Pr57^{env} is apparently derived from the amino-terminal sequence that is processed from Pr63^{env} *in vivo* during formation of Pr57^{env} (see below). The second spot (Fig. 2*C*, arrow) is present in variable amounts in the tryptic maps of both Pr63^{env} and Pr57^{env} (data not shown). When equal amounts of trypsin-digested Pr63^{env} and Pr57^{env} were mixed it was found from the peptide maps that peptides 1–6 from both proteins comigrated (Fig. 2*D*). Thus, the 63-kDa protein represents the primary translation product of the *env* gene.

It has been proposed from the DNA sequence of the Prague C strain of ASV that, during splicing to generate *env* mRNA, the reading frames of the *gag* and *env* genes are fused (ref. 6; Fig. 3*A*). According to this model, the primary translation product of the *env* gene would be a chimeric protein composed in part from *gag* gene sequences (six amino acids) while the major portion (595 amino acids) would be derived from the *env* gene (Fig. 3*B*). Pr76^{gag} from either the PrC or SRA strains contains a convenient trypsin cleavage site six amino acids downstream from the amino terminus (refs. 5, 12, and 37; Fig. 3*B*). This site corresponds to the 3′ end of the putative donor splice junction used to generate subgenomic RNAs and, as a result, trypsin digestion of either Pr76^{gag} or Pr63^{env} should release identical amino-terminal peptides. In contrast, use of the AUG codon that is part of the open reading frame of the *env* gene would produce an amino-terminal peptide of very different structure (Fig. 3*B*).

To specifically label the amino-terminal methionine, Pr76^{gag} and Pr63^{env} were synthesized *in vitro* in the presence of L-[^{35}S]methionine-tRNA$_i^{Met}$ (38, 39). The *in vitro*-synthesized proteins were isolated as described in the legend to Fig. 2. Each sample was divided into two halves: one-half was digested with L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin ("ketone-treated trypsin") and the other half was digested with α-chymotrypsin. Chymotryptic digests of amino-terminal-labeled Pr76^{gag} produced a single, major peptide spot and a minor peptide that represents digestion by contaminating trypsin activity in the α-chymotrypsin (Fig. 4*A*). This contamination was verified in mixing experiments in which amino-terminal tryptic peptide of Pr76^{gag} (Fig. 4*E*) was shown to comigrate with the minor product of chymotrypsin digestion (Fig. 4*D*). Digestion of Pr63^{env} with chymotrypsin produced a major peptide spot (Fig. 4*B*) that was clearly distinguishable from the amino-terminal chymotryptic peptide of Pr76^{gag} when the two were mixed (Fig. 4*C*). Note that as with the chymotryptic digests of Pr76^{gag} there is a minor peptide product due to contaminating trypsin activity. These results show that the polypeptides identified as Pr76^{gag} and Pr63^{env} have nonidentical amino-terminal chymotryptic peptides as predicted in Fig. 3.

Trypsin digestion of the amino-terminal-labeled proteins

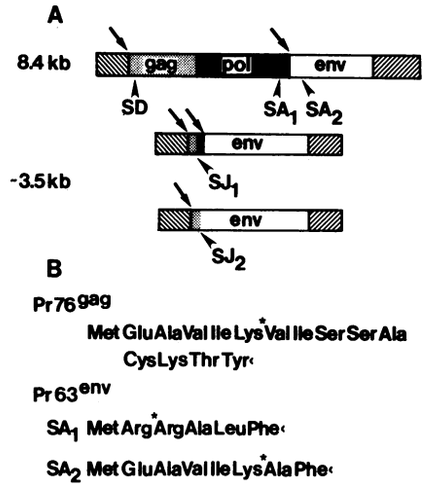

FIG. 3. Schematic representation of ASV RNA splicing and the predicted amino-terminal sequence of Pr76^{gag} and Pr63^{env}. (*A*) Genetic arrangement of the tdASV genome (8.4 kb). The positions of the donor splice junction (SD) and potential acceptor splice junctions (SA_1 and SA_2) are indicated. The positions of AUG codons present at the 5′ end of both the *gag* and *env* genes are indicated by arrows. Splicing in this system generates a 3.5-kb subgenomic RNA. The donor splice junction (SD) is joined to an acceptor splice junction (SA_1 or SA_2) at the positions marked as splice junctions (SJ_1 and SJ_2). Splicing to the upstream acceptor (SA_1) may result in use of the AUG codon at the 5′ end of the *env* gene. Alternatively, splicing to the downstream acceptor (SA_2) would fuse the *gag* and *env* genes, which would result in the synthesis of a chimeric protein. (*B*) Predicted sequences of the amino-terminal chymotryptic peptides of Pr76^{gag} and Pr63^{env}. The positions of the trypsin-sensitive sites are indicated by asterisks. The use of a splice acceptor site within the *env* gene (SA_2) will produce *gag* and *env* gene products with identical amino-terminal tryptic peptides. In contrast, use of the AUG initiation codon at the 5′ end of the *env* gene will result in nonidentical amino-terminal tryptic peptides for the *gag* and *env* gene products.

produced a single L-[^{35}S]methionine-containing peptide from either Pr76^{gag} or Pr63^{env} (Fig. 4 *E* and *F*). To confirm the identical nature of these peptides, equal amounts of each were mixed and mapped together. The result shows that these peptides, in contrast to the chymotryptic products, comigrated during both electrophoresis and ascending chromatography (Fig. 4*G*). Digestion of the mixture of chymotryptic amino-terminal peptides (Fig. 4*C*) with trypsin also resulted in a single identical product (Fig. 4*H*), confirming the position of the trypsin cleavage site as proximal to the amino terminus of both Pr63^{env} and Pr76^{gag}. It should be pointed out that the major portions of the amino-terminal peptides of Pr76^{gag} and Pr63^{env} synthesized *in vitro* (and p19 isolated from mature virions) are blocked as determined by their resistance to digestion by leucine amino peptidase (ref. 40; data not shown). Furthermore, when *in vitro* synthesis was carried out in the presence of acetonyl CoA/oxaloacetate/citrate synthase, the amino-terminal peptide was missing (peptide 7, Fig. 2*C*) from the peptide maps and was replaced by a peptide of increased positive charge that is sensitive to digestion with leucine amino peptidase. These results support the hypothesis that the amino-terminal tryptic peptides of Pr76^{gag} and Pr63^{env} are identical and that the *gag* and *env* gene sequences are fused during splicing of the *env* mRNA (Fig. 3).

Previous results have indicated that synthesis of Pr63^{env} is initiated within the 397-nucleotide leader sequence. Therefore, it is clear that these sequences must be in phase with the reading frame of the *env* gene. If the acceptor splice junction is upstream of the AUG at nucleotide 5,054, there should be two additional L-[^{35}S]methionine-labeled tryptic

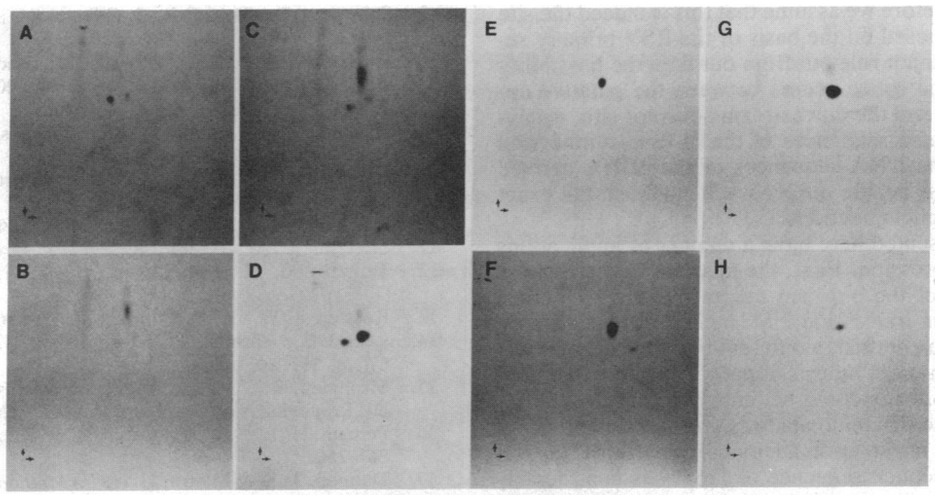

FIG. 4. Identification of the amino-terminal peptides of $Pr76^{gag}$ and $Pr63^{env}$ obtained from chymotryptic and tryptic digests. Polysomal 35S and 21S RNAs were translated in 25 μl of rabbit reticulocyte lysate in the presence of L-[^{35}S]methionine-labeled $tRNA_i^{Met}$ as described in *Materials and Methods*. The labeled proteins were excised from the dried gels and digested with α-chymotrypsin (*A–C*) or ketone-treated trypsin (*E–G*) as described in Fig. 3. (*A*) $Pr76^{gag}$ (500 cpm); (*B*) $Pr63^{env}$ (500 cpm); (*C*) $Pr76^{gag}$ together with $Pr63^{env}$ (500 cpm each); (*D*) $Pr76^{gag}$, a mixture of α-chymotrypsin- and ketone-treated trypsin-digested samples (500 cpm each); (*E*) $Pr76^{gag}$ (500 cpm); (*F*) $Pr63^{env}$ (500 cpm); (*G*) $Pr76^{gag}$ together with $Pr63^{env}$ (500 cpm each); (*H*) mixture of chymotrypsin-, ketone-treated trypsin-digested $Pr76^{gag}$ and $Pr63^{env}$. Peptide maps were prepared as described in Fig. 2.

peptides in $Pr63^{env}$ when compared with the nonglycosylated *in vivo* precursor $Pr57^{env}$. The data presented (Fig. 2) were somewhat ambiguous on this point and could not be used to exclude such a possibility.

To establish the location of the acceptor splice site, we carried out mung bean nuclease mapping of the *env* mRNA from the Prague A strain of RSV. A restriction fragment spanning the sequences encoded by the amino-terminal portion of $Pr63^{env}$ was isolated and labeled at the 5′ terminus. This fragment was then hybridized to mRNA isolated from Prague A RSV-infected cells (Fig. 5*B*). The DNA fragment protected from mung bean nuclease digestion extended four or five nucleotides beyond the consensus splice acceptor sequence (Fig. 5*C*, arrow). Apparently there is sufficient secondary structure at the termini of the hybrids to prevent the nuclease from completely digesting the labeled DNA to the end of the single-stranded regions and a single-stranded tail of four or five bases remains. Previous results from S1 mapping of the *src* mRNA have yielded similar results; i.e., the protected fragment was somewhat longer than expected (37). Thus, the *env* consensus acceptor sequence begins 24 nucleotides downstream from the AUG at nucleotide 5,054 (Fig. 4*A*). We conclude therefore that the 397 nucleotide leader sequence is spliced in phase to an acceptor site in the open reading frame encoded by the *env* gene.

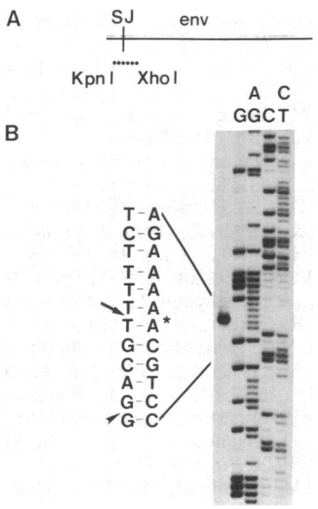

FIG. 5. Mung bean nuclease mapping of the *env* mRNA acceptor-splice junction. (*A*) The single-stranded anti-sense DNA fragment that spans the *env* mRNA acceptor splice junction was isolated from plasmid pJD100 DNA (the gift of J. T. Parsons, University of Virginia) by digestion with *Xho* I/*Kpn* I. The strands were 5′-end-labeled by using [α-^{32}P]ATP and T4 polynucleotide kinase and separated under denaturing conditions as described by Maxam and Gilbert (41). The labeled DNA was hybridized to poly(A)-containing RNA isolated from Prague A RSV-infected cells at 52°C for 3 hr, 50°C for 1 hr, and 48°C for 1 hr in 0.4 M NaCl/5 mM EDTA/0.1 M Pipes, pH 6.5/80% (vol/vol) formamide (MCB Chemical, Norwood, OH). The hybridization reaction mixture was diluted 1:20 into 50 mM NaOAc, pH 5.0/50 mM NaCl/1 mM $ZnCl_2$ and digested with mung bean nuclease as described by Patton and Chae (42). (*B*) Nuclease-resistant DNA fragments were loaded on a sequencing gel prepared as described by Maxam and Gilbert (41), next to sequence-specific reactions derived from the original *Kpn* I/*Xho* I fragment as described by Weaver and Weissman (43).

DISCUSSION

In this paper we report that $Pr76^{gag}$ and $Pr63^{env}$, the primary translation products of the ASV *gag* and *env* genes, respectively, have common amino-terminal peptides. Tryptic peptide mapping is a sensitive technique capable of separating peptides that differ by a single amino acid (44). Thus, the comigration of the amino-terminal tryptic peptides derived from $Pr63^{env}$ and $Pr76^{gag}$ during two-dimensional analysis provides strong support for their identity. The generation of nonidentical amino-terminal peptides by chymotryptic digestion of $Pr76^{gag}$ and $Pr63^{env}$ is consistent with cleavage occurring downstream from the trypsin-sensitive site—i.e., within sequences unique to either protein. Thus, our data predict that the splice donor junction used to generate subgenomic mRNA is located between the sequences that give rise to the first tryptic and chymotryptic cleavage sites of $Pr76^{gag}$. These sequences have a consensus splice donor site (10) and the data presented here are in agreement with previous S1 mapping experiments that localized the donor splice junction of the SRA RSV mRNA within the *gag* gene (11). We have localized the acceptor splice junction to the vicinity of nucleotide 5,078. This region contains a consensus splice acceptor

sequence and therefore we assume that this is indeed the site of splicing as proposed on the basis of the RSV primary sequence (5). We cannot rule out from our data the possibility that more than one splice occurs between the putative upstream donor site and the downstream acceptor site. Analysis of the amino acid sequences of the *in vitro*-synthesized proteins and of the DNA sequences of the cDNA derived from spliced virus-specific mRNAs will establish the exact positions of the splice junctions.

The results presented here have a number of implications for ASV gene expression. First, the fact that the translation initiation signals of the *gag* and *env* mRNAs are identical suggests that their translation efficiencies in infected cells are also similar. In contrast, a different initiation sequence is apparently used in the synthesis of $pp60^{src}$ (37) and therefore its translation efficiency could be quite different. Second, it is possible that the 62-amino acid-residue signal peptide derived from $Pr63^{env}$ may play an additional role in virus replication. The signal peptide sequence of $Pr63^{env}$ is considerably longer than those of most other cellular and viral glycoproteins, which have an average length 15–30 amino acids (45). A hydrophobic portion, which is characteristic of signal peptide sequences, is contained within the carboxyl-terminal half of the $Pr63^{env}$ signal peptide (6). The amino-terminal half of the signal peptide sequence, in contrast, is rich in hydrophilic and basic amino acid residues. It is known that p19, the virion structural protein derived by cleavage from the extreme amino-terminal portion of $Pr76^{gag}$, binds specifically to ASV RNA (46) and that the RNA-binding activity resides in the amino-terminal 29-amino acid CNBr cleavage fragment (47). The amino-terminal part of the signal peptide of $Pr63^{env}$ bears a strong resemblance to the amino-terminal sequence of p19 in that both have six amino acids in common at the amino terminus, as well as other homologous amino acid sequences further downstream (5, 6). This partial homology suggests the possibility that the amino-terminal part of the signal peptide may also bind to viral RNA and thus could play a role in RNA processing or translation.

Note Added in Proof. We have now determined the complete nucleotide sequence of the cDNA corresponding to the 5′ terminus of the *env* gene mRNA of tdASV and have found that there is a single splice in the mRNA from nucleotide 398 to nucleotide 5,078.

We gratefully acknowledge the expert technical assistance of Susan K. Lorenzen. We thank Steve Ellis (University of Iowa) for supplying *E. coli* extracts containing amino acyl synthetase and transformylase and C.-S. Lin and R. Stenberg for their assistance with the mung bean nuclease digestions and DNA sequence analyses. We are indebted to M. Feiss and M. F. Stinski for their critical reading of the manuscript and to Mary Smithson for typing the manuscript. This work was supported by Grant CA28051 from the National Cancer Institute. T.A.F. was a postdoctoral fellow supported by Grant IF32CA067641 from the National Cancer Institute, American Cancer Society Grant PF1921, and the Ladies Auxiliary to the Veterans of Foreign Wars.

1. Weiss, S. R., Varmus, H. E. & Bishop, J. M. (1977) *Cell* **12,** 983–992.
2. Pawson, T., Martin, G. S. & Smith, A. E. (1976) *J. Virol.* **19,** 950–967.
3. von der Helm, K. & Duesberg, P. H. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 614–618.
4. Pawson, T., Harvey, R. & Smith, A. E. (1977) *Nature (London)* **268,** 416–420.
5. Schwartz, D., Tizard, R. & Gilbert, W. (1983) *Cell* **32,** 853–869.
6. Hunter, E., Hill, E., Hardwick, M., Bhown, A., Schwartz, D. E. & Tizard, R. (1983) *J. Virol.* **46,** 920–936.
7. Sharp, P. A. (1981) *Cell* **23,** 645–646.
8. Hackett, P. B., Swanstrom, R., Varmus, H. E. & Bishop, J. M. (1982) *J. Virol.* **41,** 527–534.
9. Swanstrom, R., Varmus, H. E. & Bishop, J. M. (1982) *J. Virol.* **41,** 535–541.
10. Stoltzfus, C. M. & Snyder, P. (1975) *J. Virol.* **16,** 1161–1170.
11. Stoltzfus, C. M. & Dane, R. (1982) *J. Virol.* **42,** 918–931.
12. Stoltzfus, C. M., Dimock, K., Horikami, S. & Ficht, T. A. (1983) *J. Gen. Virol.* **64,** 2191–2202.
13. Kessler, S. W. (1975) *J. Immunol.* **115,** 1617–1624.
14. Lee, J. S., Varmus, H. E. & Bishop, J. M. (1979) *J. Biol. Chem.* **254,** 8015–8022.
15. Strohman, R. C., Moss, P. S., Micov-Eastwood, J., Spector, D., Przybyla, A. & Paterson, B. (1977) *Cell* **10,** 265–273.
16. Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* **43,** 247–256.
17. Dubnoff, J. S. & Maitra, U. (1971) *Methods Enzymol.* **20,** 248–261.
18. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
19. Bonner, W. M. & Laskey, R. A. (1974) *Eur. J. Biochem.* **46,** 83–88.
20. Chamberlain, J. P. (1979) *Anal. Biochem.* **98,** 132–135.
21. Pawson, T., Kung. T.-H. & Martin, G. S. (1981) *J. Virol.* **40,** 665–672.
22. Hayward, W. S. (1977) *J. Virol.* **24,** 47–63.
23. Brugge, J. A., Purchio, A. F. & Erikson, R. L. (1977) *Virology* **83,** 16–26.
24. Pawson, T., Mellon, P., Duesberg, P. H. & Martin, G. S. (1980) *J. Virol.* **33,** 993–1003.
25. Purchio, A. F., Jovanovich, S. & Erikson, R. L. (1980) *J. Virol.* **35,** 629–636.
26. Leamnson, R. N. & Halpern, M. S. (1976) *J. Virol.* **18,** 956–968.
27. Bolognesi, D. P., Bauer, H., Gelderblom, H. & Huper, G. (1972) *Virology* **47,** 551–566.
28. Duesberg, P. H., Martin, G. S. & Vogt, P. K. (1970) *Virology* **41,** 631–646.
29. England, J. M., Bolognesi, D. P., Dietzschold, B. & Halpern, M. S. (1977) *J. Virol.* **21,** 810–814.
30. Moelling, K. & Hayami, M. (1977) *J. Virol.* **22,** 598–607.
31. Hayman, M. V. (1978) *Virology* **85,** 475–486.
32. Buchhagen, D. L. & Hanafusa, H. (1978) *J. Virol.* **25,** 845–851.
33. Stohere, R. & Hunter, E. (1979) *J. Virol.* **32,** 412–419.
34. Diggelman, H. (1979) *J. Virol.* **30,** 799–804.
35. Shealy, D. V. & Rueckert, R. R. (1978) *J. Virol.* **26,** 380–388.
36. Klemenz, R. & Diggelmann, H. (1978) *Virology* **85,** 63–74.
37. Swanstrom, R., Parker, R. C., Varmus, H. E. & Bishop, J. M. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 2519–2523.
38. Stanley, W. M. (1972) *Anal. Biochem.* **48,** 202–216.
39. Housman, D., Jacobs-Lorena, M., Rajbhandary, J. L. & Lodish, H. F. (1970) *Nature (London)* **227,** 913–918.
40. Palmiter, R. D., Gagnon, J., Yogf, V. M., Ripley, S. & Eisenman, R. N. (1978) *Virology* **91,** 423–433.
41. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
42. Patton, J. R. & Chae, C.-B. (1983) *J. Biol. Chem.* **258,** 3991–3995.
43. Weaver, R. & Weissman, C. (1979) *Nucleic Acids Res.* **7,** 1175–1193.
44. Bond, M. W., Chiv, N. Y. & Cooperman, B. S. (1980) *Biochemistry* **19,** 94–102.
45. Austin, B. M. (1979) *FEBS Lett.* **103,** 308–313.
46. Leis, J. P., Scheible, P. & Smith, R. E. (1980) *J. Virol.* **35,** 722–731.
47. Johnson, S. P., Viegl, M., Vanaman, T. & Leis, V. (1983) *J. Virol.* **45,** 876–881.